ZooKeys 945: 1–16 (2020)
doi: 10.3897/zookeys.945.49681
https://zookeys.pensoft.net

RESEARCH ARTICLE

A peer-reviewed open-access journal
ZooKeys
Launched to accelerate biodiversity research

Characterization and comparative analysis of the complete mitochondrial genome of *Azygia hwangtsiyui* Tsin, 1933 (Digenea), the first for a member of the family Azygiidae

Yuan-An Wu[1*], Jin-Wei Gao[1*], Xiao-Fei Cheng[1], Min Xie[1], Xi-Ping Yuan[1], Dong Liu[1], Rui Song[1]

1 *Hunan Fisheries Science Institute, Changsha 410153, China*

Corresponding author: Rui Song (ryain1983@163.com)

Academic editor: D. Gibson | Received 26 December 2019 | Accepted 4 May 2020 | Published 3 July 2020

http://zoobank.org/4EE745AE-3269-4266-8CAD-FF4B67F969FB

Citation: Wu Y-A, Gao J-W, Cheng X-F, Xie M, Yuan X-P, Liu D, Song R (2020) Characterization and comparative analysis of the complete mitochondrial genome of *Azygia hwangtsiyui* Tsin, 1933 (Digenea), the first for a member of the family Azygiidae. ZooKeys 945: 1–16. https://doi.org/10.3897/zookeys.945.49681

Abstract

Azygia hwangtsiyui (Trematoda, Azygiidae), a neglected parasite of predatory fishes, is little-known in terms of its molecular epidemiology, population ecology and phylogenetic study. In the present study, the complete mitochondrial genome of *A. hwangtsiyui* was sequenced and characterized: it is a 13,973 bp circular DNA molecule and encodes 36 genes (12 protein-coding genes, 22 transfer RNA genes, two ribosomal RNA genes) as well as two non-coding regions. The A+T content of the *A. hwangtsiyui* mitogenome is 59.6% and displays a remarkable bias in nucleotide composition with a negative AT skew (–0.437) and a positive GC skew (0.408). Phylogenetic analysis based on concatenated amino acid sequences of twelve protein-coding genes reveals that *A. hwangtsiyui* is placed in a separate clade, suggesting that it has no close relationship with any other trematode family. This is the first characterization of the *A. hwangtsiyui* mitogenome, and the first reported mitogenome of the family Azygiidae. These novel datasets of the *A. hwangtsiyui* mt genome represent a meaningful resource for the development of mitochondrial markers for the identification, diagnostics, taxonomy, homology and phylogenetic relationships of trematodes.

Keywords

gene arrangement, phylogenetic analysis, Trematoda

* Contributed equally as the first authors.

Copyright *Yuan-An Wu et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

The genus *Azygia* Looss, 1899 is an endoparasitic helminth found in the stomach and intestine of freshwater feral carnivorous fish (Frolova and Shcherbina 1975). This genus includes several species complexes and its type species is *Azygia lucii* (Müller, 1776), which is a parasite of numerous, but especially esocid and percid, fishes in Europe. Many researchers have added to our knowledge of this cosmopolitan species. To date, species of *Azygia* are frequently reported from the esophagus, stomach, and intestine of a wide range of predatory fishes from Asia, Europe and North America, including China, Japan, India, Russia, Germany, and North America (Tubangui 1928; Tsin 1933; Moravec and Sey 1989; Marcogliese and Cone 1996; Besprozvannykh 2005; Jadhav et al. 2011; Pallewad et al. 2015; Womble et al. 2016; Nagasawa and Katahira 2017).

Azygia hwangtsiyui Tsin, 1933 is a member of the family Azygiidae Odhner, 1911 and is often overlooked; it is parasitic in the gastrointestinal tract of species of the family Channidae Fowler, 1934 but caused only slight clinical signs, including malnutrition and weight loss. In China, *Azygia hwangtsiyui*-infected freshwater predatory fishes have been described from Shandong, Heilongjiang, Jiangsu, Fujian, Sichuan and Hunan Provinces (Tsin 1933; Zmejev 1936; Ma 1958; Tang and Tang 1964; Kiang 1965; Chen 1973; Wang 1985; Cheng 2011). It has a mainly inland distribution and utilizes freshwater snail species (e.g. *Vivipara quadrata* (Benson, 1842)) as intermediate hosts (Tang and Tang 1964) and develops into adults in the gastrointestinal tract of predatory fish species such as *Ophiocephalus argus* Cantor, 1842 and *Channa asiatica* (Linnaeus, 1758) (Tsin 1933; Besprozvannykh 2005).

Morphology is the most commonly used method for species identification and differentiation of metazoans and is widely adopted globally by parasitologists and taxonomists. A huge disadvantage of using morphological criteria, however, is that it is difficult to identify and distinguish closely related and cryptic species. Although the family Azygiidae was erected more than a century ago, its situation, and that of several species of *Azygia*, is still controversial and uncertain. Manter (1926) pointed out that *Azygia* is the only genus in the family then presenting systematic confusion: *Azygia longa* (Leidy, 1851) in North America may be a synonym of *A. lucii* in Europe (Manter 1926), and Van Cleave and Mueller (1934) reported that *Azygia acuminata* Goldberger, 1911 and *A. longa* should be considered conspecific. Nevertheless, due to the discovery of some life histories of members of *Azygia*, *A. lucii* and *A. longa* have been recognized as two distinct species (Szidat 1932; Sillman 1953; Sillman 1962).

Mitochondrial (mt) genome and nuclear ribosomal DNA sequences are effective molecular tools for taxonomic identification, phylogeny and biogeographical research (Bernt et al. 2013; Le et al. 2019). However, only a partial cytochrome oxidase subunit 1 protein sequence (AIY67834) of *Proterometra macrostoma* Horsfall, 1933 (Azygiidae) is currently available in GenBank. None of the mitochondrial genome data have been sequenced for a member of the family Azygiidae. Therefore, we determined the complete mitochondrial genome sequence of *A. hwangtsiyui* as a basis for the future definition of strain- and species-specific markers, and for assessing mitogenomics in resolving the interrelationships of trematodes.

Materials and methods

Sampling and DNA extraction

The specimens of flatworms were isolated from the stomach of their definitive host, in this case snakehead fish (*Ophiocephalus argus* (Cantor, 1842)) obtained from east Dongting Lake in Yueyang, Hunan province, China (29°22'N, 113°06'E). *Azygia hwangtsiyui* was morphologically identified according to the original and other descriptions (Tsin 1933; Tang and Tang 1964; Zhang et al. 1999; Besprozvannykh 2005), using a stereomicroscope and a light microscope. Furthermore, single samples were confirmed molecularly as *A. hwangtsiyui* based on sequencing of 1370 bp 28S rDNA sequence. The parasites were completely washed in water, preserved in 99% ethanol, and stored at 4 °C until genomic DNA extraction. Total genomic DNA extraction was performed from an intact specimen with the TIANamp Micro DNA Kit (Tiangen Biotech, Beijing, China), according to the manufacturer's instructions.

DNA amplification and sequencing

According to conserved regions of mitochondrial genes in other available digenea mitogenomes, six partial gene fragments for cytb, nad4, nad1, 16S, 12S and cox2 were amplified using six generic primers sets HWF1/HWR1 (for cytb), HWF3/HWR3 (for nad4), HWF5/HWR5 (for nad1), HWF7/HWR7 (for 16S), HWF9/HWR9 (for 12S), and HWF11/HWR11 (for cox2), respectively. On the basis of these obtained nucleotide sequences, *A. hwangtsiyui*-specific primers were designed for amplification and sequencing of the remaining mitogenome (Suppl. material 1: Table S1). All primers were designed to produce amplicons with overlaps of approximately 100 bp. PCR reactions were performed in a 50 μl reaction solution with the ingredient of 18.5 μl ddH_2O, 25 μl 2×PCR buffer (Mg^{2+}, dNTP plus, Takara, Dalian, China), 1.5 μl of each primer (0.2–1.0 μM), 1 μl EX Taq polymerase (250U, Takara), and 2.5 μl DNA template. PCR amplification was compliant to the following amplification protocol: initial denaturation at 98 °C for 2 min, followed by 40 cycles 10 s at 98 °C, 15 s at 50 °C, 68 °C for 1 min/kb, and 10 min at 68 °C for a final extension. The amplified PCR products were purified with TIANgel Purification Kit (Tiangen Biotech, Beijing, China), and sequenced bidirectionally at Sangon Biotech (Shanghai) Co., Ltd. (Shanghai, China) based on the primer walking method using several specific primers (Suppl. material 1: Table S1).

Mitogenome annotation and analysis

According to sequence chromatograms, all raw fragments were quality-proofed using CHROMAS (https://www.technelysium.com.au) to remove ambiguity codes and low-quality bases. Whenever the quality was sub-optimal, sequencing was repeated until the amplicon is the consensus sequence. Before manual assembly of the entire mitochondrial genomic sequence, identification of all amplicons was performed by

BLASTN check (Altschul et al. 1990). The mt genome of *A. hwangtsiyui* was aligned against the mt genome sequences of other promulgated digenean mitogenomes utilizing multiple sequence alignment software MAFFT version 7.149 (Katoh and Standley 2013) to identify genetic boundary. Protein-coding genes (PCGs) were predicted with Open Reading Frame Finder (https://www.ncbi.nlm.nih.gov/orffinder/) adopting echinoderm and flatworm mitochondrial codes, and examining the nucleotide alignment against the reference mtDNA in trematode *Dicrocoelium chinense* Tang et Tang, 1978 (NC_025279.1). Whole tRNAs were inferred with the detection results of ARWEN (Laslett and Canback 2008) and MITOS web server (Bernt et al. 2013). Two rRNA (rrnL and rrnS) were founding by comparison with those of published fluke mitogenomes. Codon usage and relative synonymous codon usage (RSCU) for 12 PCGs of *A. hwangtsiyui* were computed by PHYLOSUITE (Zhang et al. 2020), and its operation results were imported into GGPLOT2 program (Wickham 2016) to make figures of the RSCU. Tandem repeats in the non-coding regions were determined with Tandem Repeats Finder software version 4.09 (Benson 1999), and the prediction of their secondary structures were performed by the MFOLD web server (Zuker 2003). The annular diagram of *A. hwangtsiyui* mitogenome was plotted with mitochondrial genome data visualization tool MTVIZ (http://pacosy.informatik.uni-leipzig.de/mtviz/mtviz).

Phylogenetic analysis

For phylogenetic analyses, we utilized translated and concatenated amino acid sequences of twelve protein-coding genes for 49 Platyhelminthes including *A. hwangtsiyui* mitogenome determined in this study. Two tapeworm species, *Cloacotaenia megalops* (Nitzsch in Creplin, 1829) (NC_032295.1) and *Dibothriocephalus latus* (Linnaeus, 1758) (NC_008945.1) were included as outgroup taxa representing two different families. Species information including systematic positions and GenBank accession numbers is provided in Suppl. material 2: Table S2. The PHYLOSUITE program was used to extract twelve PCGs from the GenBank files, export fasta files with translated amino acid datasets, and align datasets in bulk using integrated applet MAFFT with normal-alignment mode. Phylogenetic analyses were performed using Bayesian Inference (BI) and Maximum Likelihood (ML) methods. Assessment of the best-fit evolutionary model for dataset was conducted via ModelGenerator v0.8527 (Keane et al. 2006). BI in MrBayes version 3.2.6 (Ronquist et al. 2012) was carried out under the MtRev matrix of amino acid substitution, and was analyzed with 1×10^7 metropolis-coupled Monte Carlo Markov Chain (MCMC) generations. Two independent runs with four simultaneous MCMC chains (one cold and three heated chains) were conducted for 1×10^7 million generations, sampling every 10,000 generations and discarding the initial 25% generations as burn-in. ML analysis in PHYLOSUITE was performed using MtART+I+G matrix with 1000 bootstrap replicates.

Results and discussion

General traits of the *Azygia hwangtsiyui* mitogenome

The entire *A. hwangtsiyui* mtDNA is 13,973 bp in length (GenBank accession number: MN844889) and comprised of 12 protein-coding genes (cox1-3, nad1-6, nad4L, cytb, and atp6), 22 tRNA genes, two rRNA genes (rrnL and rrnS), and two non-coding regions. The 12 protein-coding gene order arrangement is cox3-cytb-nad4L-nad4-atp6-nad2-nad1-nad3-cox1-cox2-nad6-nad5 (Fig. 1), which is identical to those of *Clinostomum complanatum* (Rudolphi, 1814), *Echinostoma hortense* Asada, 1926, and some species of the Fasciolidae (*Fasciola hepatica* Linneuus, 1758, *Fasciola gigantica* Cobbold, 1856, and *Fasciola* sp. GHL-2014) (Liu et al. 2014a; Chen et al. 2016; Liu et al. 2016); the gene atp8 is similarly missing, as usual in trematode species. All genes are transcribed in the anticlockwise direction and encoded by H strand (Table 1), which is in accordance with other digeneans. The mt genome of *A. hwangtsiyui* has 22 intergenic spacers ranging from 1 to 15 bp and contains two overlapping nucleotides ranging from 1 to 40 bp (Table 1). Noteworthily, a 40 bp overlap between the nad4 and nad4L genes exists in the *A. hwangtsiyui* mitogenome, which is consistent with most helminths such as *Eurytrema pancreaticum* Janson, 1889 (Chang et al. 2016), *Hypoderaeum conoideum* (Bloch, 1782) (Yang et al. 2015), but shorter than that of *Schistosoma mekongi* Voge, Bruckner & Bruce, 1978 (64 bp; Littlewood et al. 2006). The nucleotide contents of T, C, A, G, in *A. hwangtsiyui* mitogenome are 42.8%, 12.0%, 16.8%, and 28.5%, respectively (Table 2). The whole A+T content of the mitogenome is 59.6%, which was markedly biased toward T over A (AT skew: –0.437), and G over C (GC skew: 0.408).

Protein-coding genes and non-coding regions

A total of 3364 amino acids was encoded by the *A. hwangtsiyui* mtDNA. The full scale of 12 concatenated protein-coding genes was 10126 bp, composed of 45.2% T, 11.5% C, 14.7% A, and 28.6% G. Average A+T content of concatenated 12 protein-coding genes was 59.9%, varying from 57.7% (cox2) to 64.8% (nad4L) (Table 2 and Suppl. material 2: Table S2). All 12 protein-coding genes of *A. hwangtsiyui* mt genome have a lower A+T percentage than those of *Trichobilharzia szidati* Neuhaus, 1952, *Calicophoron microbothrioides* Price & McIntosh, 1944, and some members of the Schistosomatidae Poche, 1907, but possess a higher A+T percentage than those of *Metagonimus yokogawai* Katsurada, 1912, and *Paragonimus westermani* Kerbert, 1878 (Suppl. material 3: Table S3) (Lee et al. 1987; Littlewood et al. 2006; Biswal et al. 2014; Semyenova et al. 2017; Oey et al. 2019). There is an obvious bias towards T over A (AT skew = –0.509), and G over C (GC skew = 0.425), and the coding strand is enriched with T and poor with A and especially C. For *A. hwangtsiyui*, the length of

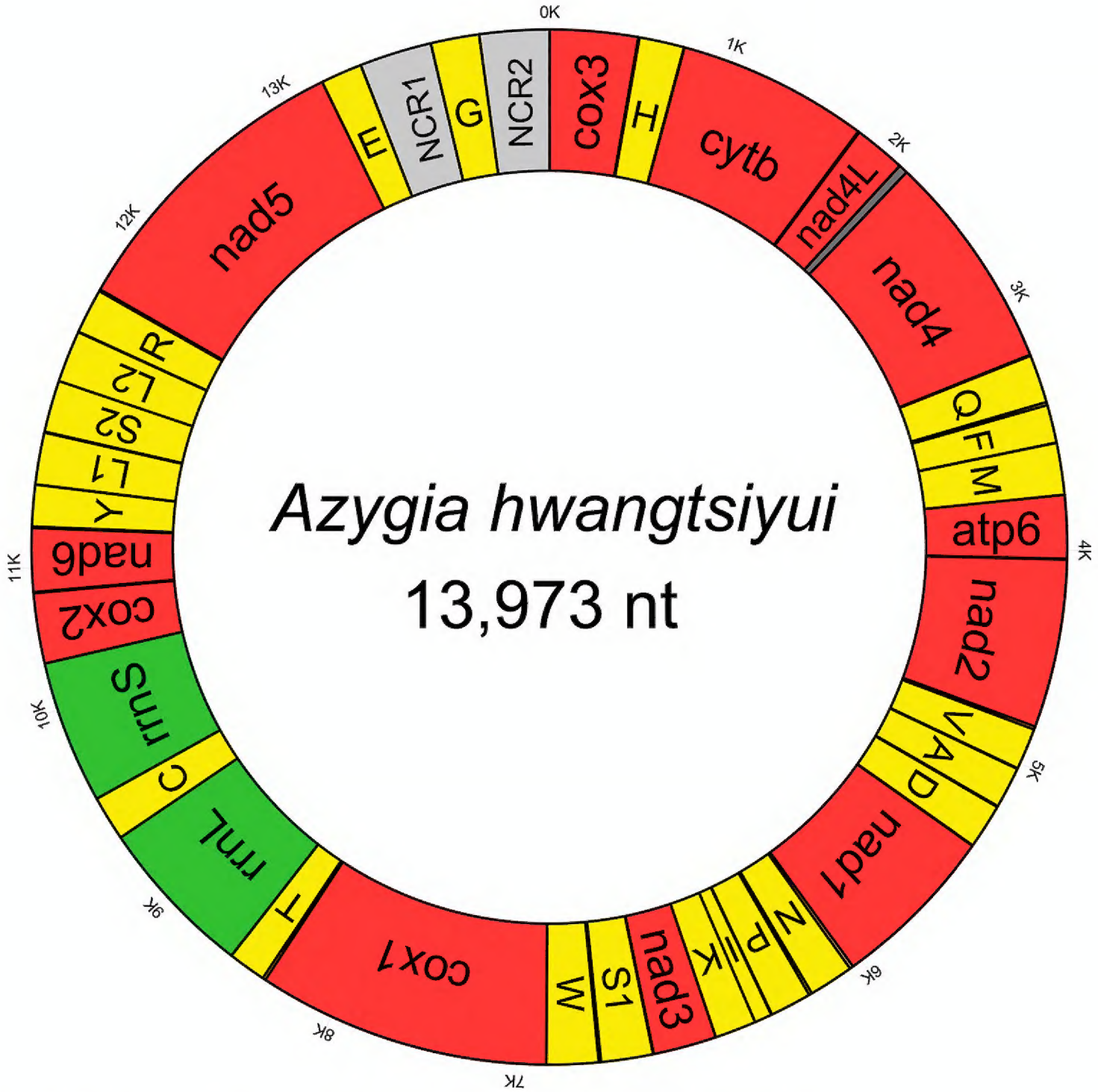

Figure 1. An annular diagram of the *Azygia hwangtsiyui* mitochondrial genome.

protein-coding genes was followed in the order: nad5 (1600 bp) > cox1 (1554 bp) > nad4 (1272 bp) > cytb (1110 bp) > nad1 (906 bp) > nad2 (861 bp) > cox3 (660 bp) > cox2 (582 bp) > atp6 (513 bp) > nad6 (444 bp) > nad3 (360 bp) > nad4L (264 bp). There are two non-coding regions (NCR1 and NCR2) in *A. hwangtsiyui* mitogenome, while the mt genome of *Paragonimus heterotremus* Chen et Hsia (1964), *C. complanatum*, *Fascioloides magna* (Bassi, 1875), and *T. szidati* have a single non-coding region (Chen et al. 2016; Ma et al. 2016; Semyenova et al. 2017; Qian et al. 2018). NCR1 and NCR2 of the *A. hwangtsiyui* mitogenome is partitioned into two parts by trnG, and accompanied by 70.5% and 57.6% A+T content, respectively. NCR1 and NCR2 have similar chemical base counts, 315 bp and 317 bp in size, respectively. While the NCR1 lacks distinguishing features and any tandem repeats, the NCR2 contains two typical tandem repeats, and each of tandem repeats sequence (120 bp) forms a hairpin-

Table 1. The organization of the mitochondrial genome of *Azygia hwangtsiyui*.

Gene	Position		Size	Intergenic nucleotides	Codon		Anti-codon	Strand
	From	To			Start	Stop		
cox3	1	660	660	–	ATG	TAG	–	H
trnH	666	729	64	+5	–	–	GTG	H
cytb	732	1841	1110	+2	ATG	TAG	–	H
nad4L	1848	2108	261	+6	ATG	TAG	–	H
nad4	2069	3340	1272	–40	ATG	TAG	–	H
trnQ	3345	3409	65	+4	–	–	TTG	H
trnF	3423	3488	66	+13	–	–	GAA	H
trnM	3490	3555	66	+1	–	–	CAT	H
atp6	3556	4068	513	–	ATG	TAG	–	H
nad2	4072	4932	861	+3	GTG	TAG	–	H
trnV	4946	5009	64	+13	–	–	TAC	H
trnA	5013	5076	64	+3	–	–	TGC	H
trnD	5081	5146	66	+4	–	–	GTC	H
nad1	5149	6054	906	+2	GTG	TAG	–	H
trnN	6070	6134	65	+15	–	–	GTT	H
trnP	6148	6212	65	+13	–	–	TGG	H
trnI	6216	6279	64	+3	–	–	GAT	H
trnK	6280	6348	69	–	–	–	CTT	H
nad3	6349	6708	360	–	ATG	TAA	–	H
trnS1	6712	6770	59	+3	–	–	GCT	H
trnW	6781	6842	62	+10	–	–	TCA	H
cox1	6843	8396	1554	–	TTG	TAG	–	H
trnT	8410	8474	65	+13	–	–	TGT	H
rrnL	8475	9449	975	–	–	–	–	H
trnC	9450	9506	57	–	–	–	GCA	H
rrnS	9507	10246	740	–		–	–	H
cox2	10247	10828	582	–	GTG	TAA	–	H
nad6	10834	11277	444	+5	GTG	TAG	–	H
trnY	11284	11352	69	+6	–	–	GTA	H
trnL1	11352	11416	65	–1	–	–	TAG	H
trnS2	11421	11490	70	+4	–	–	TGA	H
trnL2	11491	11555	65	–	–	–	TAA	H
trnR	11558	11617	60	+2	–	–	TCG	H
nad5	11626	13225	1600	+8	GTG	T	–	H
trnE	13226	13288	63	–	–	–	TTC	H
trnG	13604	13669	66	–	–	–	TCC	H

like secondary structure including a whole set of stems and loops (Suppl. material 4: Figure S1). Although tandem repeats are a segment of function-deficiency mitochondrial genome sequences, its hairpin-like secondary structures are widely perceived as regulating the replication and transcription of mitochondrial genome.

Codon usage, transfer RNAs, and ribosomal RNAs

For the *A. hwangtsiyui* mitogenome, codon ends in G or T were more continual than those ending in A or C. The most frequently used start codon in protein-coding genes was ATG (for six PCGs), secondly was GTG (for five PCGs), which resembles that of the most frequent extrapolated start codons for mitogenome protein-encoding genes

Table 2. Nucleotide contents of genes and the non–coding region within the mitochondrial genome of *Azygia hwangtsiyui*.

Regions	Size (bp)	T	C	A	G	AT (%)	GC (%)	AT skew	GC skew
atp6	513	50.5	11.3	12.9	25.3	63.4	36.6	–0.594	0.383
cox1	1554	44.0	13.0	15.9	27.1	59.9	40.1	–0.469	0.352
cox2	582	39.7	13.1	18.0	29.2	57.7	42.3	–0.375	0.382
cox3	660	44.8	11.4	15.9	27.9	60.7	39.3	–0.476	0.421
cytb	1110	44.4	13.5	15.8	26.3	60.2	39.8	–0.476	0.321
nad1	906	44.0	10.0	16.3	29.6	60.3	39.6	–0.459	0.493
nad2	861	46.2	11.3	11.8	30.7	58.0	42.0	–0.592	0.463
nad3	360	45.8	8.1	14.2	31.9	60.0	40.0	–0.528	0.597
nad4	1272	45.7	12.8	12.3	29.2	58.0	42.0	–0.577	0.391
nad4L	264	47.0	8.0	17.8	27.3	64.8	35.3	–0.450	0.548
nad5	1600	46.4	9.5	14.1	30.0	60.5	39.5	–0.535	0.519
nad6	444	45.9	11.9	14.2	27.9	60.1	39.8	–0.528	0.401
rrnL	975	37.4	14.1	22.5	26.1	59.9	40.2	–0.250	0.299
rrnS	740	36.6	14.5	21.6	27.3	58.2	41.8	–0.258	0.307
tRNAs	1349	35.3	14.5	21.7	28.5	57.0	43.0	–0.238	0.328
rRNAs	1715	37.1	14.2	22.1	26.6	59.2	40.8	–0.253	0.303
PCGs	10126	45.2	11.5	14.7	28.6	59.9	40.1	–0.509	0.425
Genome	13973	42.8	12.0	16.8	28.5	59.6	40.5	–0.437	0.408

of digenean species (Chen et al. 2016). The least-used start codon was TTG (only one PCGs). Likewise, the most-used terminal codon was TAG (for nine PCGs), followed by TAA (for two PCGs). Only one of 12 protein-coding sequence (nad5) was terminated with abbreviated T stop codon (Table 1). Although incomplete stop codons (T or TA) frequently occur in cestodes and nematodes, they were rarely presented in flukes other than *D. chinensis*, *Dicrocoelium dendriticum* (Rudolphi, 1819), and *Postharmostomum commutatum* (Dietz, 1858) (Liu et al. 2014b; Fu et al. 2019). The codon UUU (Phe, 10.17%), UUG (Leu, 8.17%), and GUU (Val, 7.19%) were the most frequently occurring codons in protein-coding genes. Leucine, valine, and phenylalanine are the most-used amino acids, with frequency of 15.96%, 13.38%, and 11.09%, respectively. The least-used codons were CGA (Arg, 0.06%) and GAC (Asp, 0.15%), and the least frequent utilized amino acid was glutamine (1.01%) (Suppl. material 5: Figure S2). As most of digenea mitogenome sequences, the mitogenome of *A. hwangtsiyui* possessed 22 commonly found tRNAs, with the exception of that of *P. westermani* Korean isolate (23 tRNAs), and *P. westermani* Indian isolate (24 tRNAs) (Biswal et al. 2014). In *A. hwangtsiyui*, tRNA-Gly (trnG) is located between NCR1 and NCR2 (Fig. 1). The size of ribosomal RNA genes (rrnL and rrnS) in mitochondrial DNA of *A. hwangtsiyui* are 975 bp and 740 bp, respectively (Table 2). The upstream and downstream of rrnL and rrnS are cascaded with trnT and cox2 genes, respectively, and are detached from each other by trnC, as in all reported platyhelminths to date (Littlewood et al. 2006, Le et al. 2016).

Gene arrangement

Comparative analysis of gene arrangement among 47 selected digenean taxa, two gene blocks (cox1-trnT-rrnL-trnC-rrnS-cox2-nad6 and cytb-nad4L-nad4-trnQ) are shared

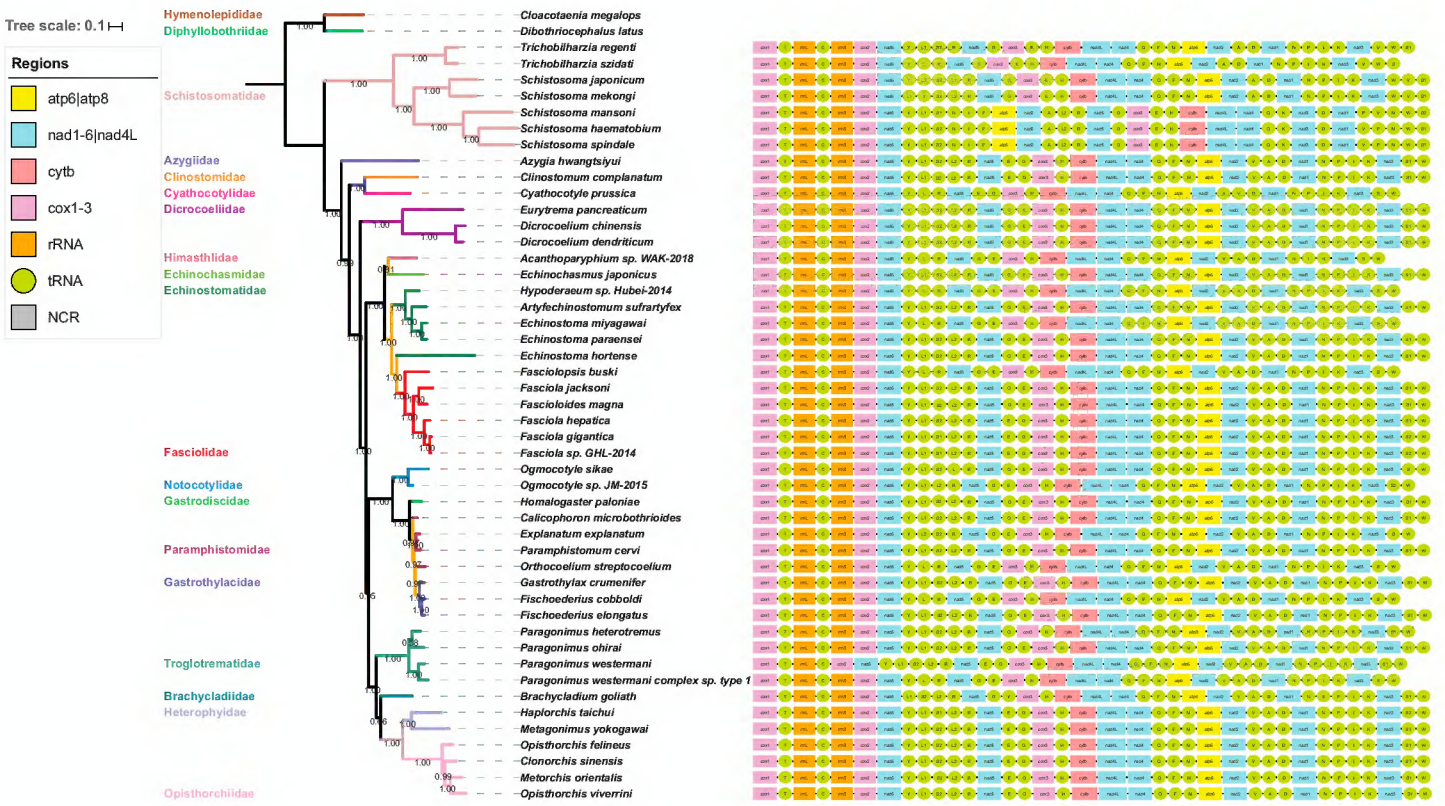

Figure 2. Phylogenetic relationships and gene arrangement of *Azygia hwangtsiyui* with other selected digeneas based on translated mitochondrial proteins. The concatenated amino-acid sequence datasets of the 12 protein-coding genes were analyzed by Bayesian Inference (BI) and Maximum Likelihood (ML), utilizing *Cloacotaenia megalops* (NC_032295.1) and *Dibothriocephalus latus* (NC_008945.1) as the outgroups. Both ML and BI analyses constructed identical tree topologies.

by all selected taxa (Fig. 2). Disregarding *P. heterotremus* and members of the family Schistosomatidae and Fasciolidae Raillet, 1895, the gene order of the remaining digenea taxa is virtually identical with the exception of the translocation of trnE and trnG among the remaining members of selected digenea representatives in family level. Intriguingly, there is the translocation of trnE and trnG within different species of family Fasciolidae. The translocations of three tRNAs (trnS1, trnS2 and trnS) can be discovered even between taxa of the same subgroup. Gene order of the *Brachycladium goliath* (Van Beneden, 1858) mt genome (the only representative of family Brachycladiidae Faust, 1929) is nearly same as that of *P. westermani* (Troglotrematidae Ward, 1918) except for the relocations of trnY between trnG and cox3, and trnE to the position between trnN and trnP. The groups of Schistosomatidae show a massive gene reorganization of protein-coding genes and tRNAs compared with other sequenced digenea mitogenome, which is in accord with previous finding reported by Littlewood et al. (2006).

Mitogenome-derived phylogeny

To assess phylogenetic relationships among available flatworms, we utilized concatenated amino acid sequence dataset representing 12 protein-coding genes of *A. hwangtsiyui*, 46 other digenean representatives, and two tapeworm species (*C. megalops* and *D. latus*) for analyzing molecular-based phylogeny. In this study, the topological structure is divided into two large clades: one consists of seven members of the family Schis-

tosomatidae; and the other clade comprises 40 members from 16 families including the family Azygiidae (*A. hwangtsiyui*) (Fig. 2). The topological structure shows that *A. hwangtsiyui* (Azygiidae) is identified as the most basal lineage of the Digenea, but separated from *C. complanatum* (Clinostomidae Lühe, 1901), and *Cyathocotyle prussica* Mühling, 1896 (Cyathocotylidae Poche, 1926). Phylogenetic analyses of all complete digenea mtDNAs confirmed taxonomic and previous phylogenetic assessments (Olson et al. 2003; Kostadinova and PéRez-del-Olmo 2014; Fu et al. 2019). The intricate structure and varying content of the family Azygiidae still awaits investigation of relationships based on a much wider taxon sampling and more mitogenome datasets.

Acknowledgments

This work was supported by the Earmarked Fund for China Agriculture Research System (CARS-45) and Important Research Project of Hunan Provincial Science and Technology Department (2016NK2176).

References

Altschul SF, Gish W, Miller W, Myer EW, Lipman D (1990) Basic local alignment search tool (blast). Journal of Molecular Biology 215(3): 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research 27(2): 573–580. https://doi.org/10.1093/nar/27.2.573

Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution 69(2): 313–319. https://doi.org/10.1016/j.ympev.2012.08.023

Besprozvannykh VV (2005) Life cycles of the trematode species *Azygia hwangtsiytii* and *A. robusta* (Azygiidae) in Primorsky Territory. Parazitologiia 39(4): 278–284.

Biswal DK, Chatterjee A, Bhattacharya A, Tandon V (2014) The mitochondrial genome of *Paragonimus westermani* (Kerbert, 1878), the Indian isolate of the lung fluke representative of the family Paragonimidae (Trematoda). Peerj 2: e484. https://doi.org/10.7717/peerj.484

Chang QC, Liu GH, Gao JF, Zheng X, Zhang Y, Duan H, Yue DM, Fu X, Su X, Gao Y (2016) Sequencing and characterization of the complete mitochondrial genome from the pancreatic fluke *Eurytrema pancreaticum* (trematoda: Dicrocoeliidae). Gene 576(1): 160–165. https://doi.org/10.1016/j.gene.2015.09.081

Chen CL (1973) An Illustrated Guide to the Fish Diseases and causative Pathogenic Fauna and Flora in the Hubei Province. Science Press, Beijing, 271 pp.

Chen L, Feng Y, Chen HM, Wang LX, Feng HL, Yang X, Mughal MN, Fang R (2016) Complete mitochondrial genome analysis of *Clinostomum complanatum* and its comparison with selected digeneans. Parasitology Research 115(8): 3249–3256. https://doi.org/10.1007/s00436-016-5102-6

Cheng YD (2011) Human and Animal Parasitic Helminthes in Hunan, China. Hunan Science and Technology Press, Changsha, 762 pp.

Frolova EN, Shcherbina TV (1975) New species in genus *Azygia* Looss, 1899 (Trematoda, Azygiidae). Parazitologiia 9(6): 489–493.

Fu YT, Jin YC, Liu GH (2019) The complete mitochondrial genome of the caecal fluke of poultry, *Postharmostomum commutatum*, as the first representative from the superfamily Brachylaimoidea. Frontiers in Genetics 10: 1–13. https://doi.org/10.3389/fgene.2019.01037

Jadhav S, Anarse S, Borde SN (2011) Incidence of helminth parasites in fresh water fishes from Sina Kolegoan Dam, Dist. Osmanabad (MS) India. Recent Research in Science and Technology 3(3): 49–50.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution 30(4): 772–780. https://doi.org/10.1093/molbev/mst010

Keane TM, Creevey CJ, Pentony MM, Naughton TJ, McInerney JO (2006) Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. BMC Evolutionary Biology 6: 1–17. https://doi.org/10.1186/1471-2148-6-29

Kiang TC (1965) Digenetic trematodes of the *Ophiocephalus argus*, *Cyprinus carpio*, and *Parasilurus asotus* in Chengtu I. Journal of Sichuan University (Natural Science Edition) 1(1): 115–122.

Kostadinova A, PéRez-del-Olmo A (2014) The systematics of the Trematoda. Advances in Experimental Medicine and Biology 766: 21–44. https://doi.org/10.1007/978-1-4939-0915-5_2

Laslett D, Canback B (2008) ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics 24(2): 172–175. https://doi.org/10.1093/bioinformatics/btm573

Le TH, Nguyen NTB, Nguyen KT, Doan HTT, Dung DT, Blair D (2016) A complete mitochondrial genome from *Echinochasmus japonicus* supports the elevation of Echinochasminae Odhner, 1910 to family rank (Trematoda: Platyhelminthes). Infection Genetics and Evolution 45: 369–377. https://doi.org/10.1016/j.meegid.2016.09.024

Lee CC, Sheikhomar AR, Jones A (1987) *Calicophoron microbothrioides* (Price and McIntosh, 1944) (Paramphistomidae: Paramphistominae) in Malaysian sambar deer (*Cervus unicolor*). New Zealand Veterinary Journal 35(11): 190–191. https://doi.org/10.1080/00480169./1987.35442

Littlewood DTJ, Lockyer AE, Webster BL, Johnston DA, Le TH (2006) The complete mitochondrial genomes of *Schistosoma haematobium* and *Schistosoma spindale* and the evolutionary history of mitochondrial genome changes among parasitic flatworms. Molecular Phylogenetics and Evolution 39(2): 452–467. https://doi.org/10.1016/j.ympev.2005.12.012

Liu GH, Gasser RB, Young ND, Song HQ, Ai L, Zhu XQ (2014a) Complete mitochondrial genomes of the 'intermediate form' of *Fasciola* and *Fasciola gigantica*, and their comparison with *F. hepatica*. Parasites & Vectors 7(1): 1–10. https://doi.org/10.1186/1756-3305-7-150

Liu GH, Yan HB, Otranto D, Wang XY, Zhao GH, Jia WZ, Zhu XQ (2014b) *Dicrocoelium chinensis* and *Dicrocoelium dendriticum* (Trematoda: Digenea) are distinct lancet fluke species based on mitochondrial and nuclear ribosomal DNA sequences. Molecular Phylogenetics and Evolution 79: 325–331. https://doi.org/10.1016/j.ympev.2014.07.002

Liu ZX, Zhang Y, Liu YT, Chang QC, Su X, Fu X, Yue DM, Gao Y, Wang CR (2016) Complete mitochondrial genome of *Echinostoma hortense* (Digenea: Echinostomatidae). Korean Journal of Parasitology 54(2): 173–179. https://doi.org/10.3347/kjp.2016.54.2.173

Lockyer AE, Olson PD, Ostergaard P, Rollinson D, Johnston DA, Attwood SW, Southgate VR, Horak P, Snyder SD, Le TH, Agatsuma T, McManus DP, Carmichael AC, Naem S, Littlewood DTJ (2003) The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. Parasitology 126(3): 203–224. https://doi.org/10.1017/S0031182002002792

Looss A (1899) Weitere Beiträge zur Kenntnis der trematoden-fauna Äezyptens, zugleich versuch einer natürlichen gliederung des genus *Distomum* Retzius. Zoologische Jahrbuecher Abteilung fuer Systematik Oekologie und Geographie der Tiere 12: 521–784. https://doi.org/10.5962/bhl.part.2037

Ma CL (1958) Worm parasites from Taihu fishes: digenetic trematodes IV. *Azygia* (Azygiidae). Journal of East China Normal University (Natural Science) 1: 26–28.

Ma J, He JJ, Liu GH, Leontovyc R, Kasny M, Zhu XQ (2016) Complete mitochondrial genome of the giant liver fluke *Fascioloides magna* (Digenea: Fasciolidae) and its comparison with selected trematodes. Parasites & Vectors 9: 1–7. https://doi.org/10.1186/s13071-016-1699-7

Manter HW (1926) Some North American fish trematodes. Illinois Biology Monographs 10: 1–138. https://doi.org/10.5962/bhl.title.50240

Marcogliese DJ, Cone DK (1996) On the distribution and abundance of eel parasites in Nova Scotia: influence of pH. Journal of Parasitology 82(3): 389–399. https://doi.org/10.2307/3284074

Moravec F, Sey O (1989) Some trematodes of freshwater fishes from North Vietnam with a list of recorded endohelminths by fish hosts. Folia Parasitologica 36: 243–262.

Nagasawa K, Katahira H (2017) A revised and updated checklist of the parasites of eels (*Anguilla* spp.) (Anguilliformes: Anguillidae) in Japan (1915–2017). Biosphere Science 56: 33–69.

Oey H, Zakrzewski M, Narain K, Devi KR, Agatsuma T, Nawaratna S, Gobert GN, Jones MK, Ragan MA, McManus DP, Krause L (2019) Whole-genome sequence of the oriental lung fluke *Paragonimus westermani*. Gigascience 8: 1–8. https://doi.org/10.1093/gigascience/giy146

Olson PD, Cribb TH, Tkach VV, Bray RA, Littlewood DTJ (2003) Phylogeny and classification of the Digenea (Platyhelminthes: Trematoda). International Journal for Parasitology 33(7): 733–755. https://doi.org/10.1016/S0020-7519(03)00049-3

Pallewad S, Nanware SS, Bhure DB (2015) Incidence of infection of trematode genus *Azygia* Looss, 1899 parasitic in freshwater fish *Channa punctatus*. Biolife 3: 187–191.

Qian L, Zhou P, Li WL, Wang HP, Miao TJ, Hu LG (2018) Characterization of the complete mitochondrial genome of the lung fluke, *Paragonimus heterotremus*. Mitochondrial DNA Part B-Resources 3(2): 560–561. https://doi.org/10.1080/23802359.2018.1462119

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Systematic Biology 61(2): 539–542. https://doi.org/10.1093/sysbio/sys029

Semyenova S, Chrisanfova G, Mozharovskaya L, Guliaev A, Ryskov A (2017) The complete mitochondrial genome of the causative agent of the human cercarial dermatitis, the visceral

bird schistosome species *Trichobilharzia szidati* (platyhelminthes: Trematoda: Schistosomatidae). Mitochondrial DNA Part B-Resources 2(2): 469–470. https://doi.org/10.1080/23802359.2017.1347833

Sillman EI (1953) The life history of *Azygia longa* (Leidy, 1851) (Trematoda: Azygiidae). Journal of Parasitology 39: 1–15.

Sillman EI (1962) The life history of *Azygia longa* (Leidy, 1851) (Trematoda: Digenea), and notes on *A. acuminata* Goldberger 1911. Transactions of the American Microscopical Society 81(1): 43–65. https://doi.org/10.2307/3223942

Szidat L (1932) Über cysticerke Riesen cercarien, insbesondere *Cercaria mirabilis* M. Braun and *Cercaria splendens* n. sp., und ihre Entwicklung im Magen von Raubfischen zu Trematoden der Gattung *Azygia* Looss. Zeitschrift Für Parasitenkunde 4(3): 477–505. https://doi.org/10.1007/BF02119996

Tang CC, Tang CT (1964) Studies on the life history of *Azygia anguillae* Ozaki 1924 and *Azygia hwangtsiyüi* Tsin 1933, with a discussion on the systematics of the genus. Acta Parasitologica Sinica 1: 137–152.

Tsin SM (1933) A new species of trematode, *Azygia hwangtsiyüi*, from *Ophiocephalus argus*. Journal of Science National University of Shantung 1: 104–116.

Tubangui MA (1928) Trematode parasites of Philippine vertebrates. Philippine Journal of Science 36: 351–371.

Van Cleave HJ, Mueller JF (1934) Parasites of Oneida Lake fishes, part III: A biological and ecological survey of the worm parasites. part IV: Additional notes on parasites of Oneida Lake fishes, including descriptions of new species. Roosevelt Wild Life Annals 3: 161–373. https://doi.org/10.2307/3222099

Wang XY (1985) Catalogue of helminthes of fishes in Poyang Lake. Jiangxi Science 3(1): 34–48.

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York, 260 pp. https://doi.org/10.1002/wics.147

Womble MR, Orélis-Ribeiro R, Bullard SAJPI (2016) New species of *Proterometra* (Digenea: Azygiidae) and its life cycle in the Chickasawhay River, Mississippi, USA, with supplemental observations of *Proterometra autraini*. Parasitology International 65(1): 31–43. https://doi.org/10.1016/j.parint.2015.09.005

Yang X, Gasser RB, Koehler AV, Wang LX, Zhu KX, Chen L, Feng HL, Hu M, Fang R (2015) Mitochondrial genome of *Hypoderaeum conoideum* – comparison with selected trematodes. Parasites & Vectors 8: 1–8. https://doi.org/10.1186/s13071-015-0720-x

Zhang D, Gao FL, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources 20(1): 348–355. https://doi.org/10.1111/1755-0998.13096

Zhang JY, Qiu ZZ, Ding X J (1999) Parasites and Parasitic Diseases of Fishes. Science Press, Beijing, 735 pp.

Zmejev GJ (1936) Les trematodes et les cestodes des poisons de l'Amour. Parazitologicheski Sbornik 6: 405–436.

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31(13): 3406–3415. https://doi.org/10.1093/nar/gkg595

Supplementary material 1

Table S1

Authors: Yuan-An Wu, Jin-Wei Gao, Xiao-Fei Cheng, Min Xie, Xi-Ping Yuan, Dong Liu, Rui Song

Data type: molecular data

Explanation note: Primers for amplification and sequencing mitochondrial genome of *Azygia hwangtsiyui*.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.945.49681.suppl1

Supplementary material 2

Table S2

Authors: Yuan-An Wu, Jin-Wei Gao, Xiao-Fei Cheng, Min Xie, Xi-Ping Yuan, Dong Liu, Rui Song

Data type: molecular data

Explanation note: Information of the Digenea and the outgroups for which complete mitogenomes are available in GenBank.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.945.49681.suppl2

Supplementary material 3

Table S3

Authors: Yuan-An Wu, Jin-Wei Gao, Xiao-Fei Cheng, Min Xie, Xi-Ping Yuan, Dong Liu, Rui Song

Data type: molecular data

Explanation note: A+T content (%) for 12 protein-coding genes in the available 49 Platyhelminthes mitogenomes.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.945.49681.suppl3

Supplementary material 4

Figure S1

Authors: Yuan-An Wu, Jin-Wei Gao, Xiao-Fei Cheng, Min Xie, Xi-Ping Yuan, Dong Liu, Rui Song

Data type: molecular data

Explanation note: Secondary structure of tandem repeats in non-coding region 2 (NCR2) of *Azygia hwangtsiyui* mitochondrial genome.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.945.49681.suppl4

Supplementary material 5

Figure S2

Authors: Yuan-An Wu, Jin-Wei Gao, Xiao-Fei Cheng, Min Xie, Xi-Ping Yuan, Dong Liu, Rui Song

Data type: molecular data

Explanation note: Relative synonymous codon usage (RSCU) of *Azygia hwangtsiyui* mitochondrial genome.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/zookeys.945.49681.suppl5